Sex Abuse (2006) 18:383–400
DOI 10.1007/s11194-006-9024-1

ORIGINAL ARTICLE

Assessing Sexual Arousal with Adolescent Males Who Have Offended Sexually: Self-Report and Unobtrusively Measured Viewing Time

James R. Worling

Published online: 2 December 2006

Abstract Sexual arousal was assessed using three approaches: the Affinity (Version. 1.0) computerized assessment of unobtrusively measured viewing time (VT), Affinity self-report ratings of sexual attractiveness, and a self-report sexual arousal graphing procedure. Data were collected from 78 males, aged 12–18 ($M = 15.09$; $SD = 1.62$), who acknowledged their sexual assaults. The pattern of responses to all three assessment techniques was remarkably similar, with maximal sexual interest demonstrated and reported for adolescent and adult females. Both self-report procedures could significantly distinguish those adolescents who assaulted a child from those who assaulted peers or adults. The self-report procedures could also significantly discriminate those adolescents with male child victims. The Affinity VT approach significantly differentiated those adolescents who assaulted male children from those who assaulted other individuals. No assessment technique could accurately identify those adolescents with exclusively female child victims. Overall, the results suggest that structured, self-report data regarding sexual interests can be useful in the assessment of adolescents who have offended sexually.

Keywords Sexual arousal · Sexual interests · Self-report · Viewing time · Adolescents · Sexual offending

J. R. Worling (✉)
Sexual Abuse: Family Education & Treatment (SAFE-T) Program, Thistletown Regional Centre, Ontario Ministry of Children and Youth Services, SAFE-T Program, 51 Panorama Crt., Toronto, Ontario, Canada, M9V 4L8
e-mail: jworling@ican.net

J. R. Worling
Department of Human Development & Applied Psychology, OISE/University of Toronto, Toronto, Canada

Introduction

Much of the treatment that was provided in the 1980's and 1990's to adolescents who offended sexually was predicated on the notion that deviant sexual interests played a critical role in the sexual assaults. For example, adolescents were often asked to track their sexual interests in violence and/or prepubescent children for lengthy time periods by completing deviant fantasy logs and charts and to participate in punishment-based procedures, such as covert sensitization, to reduce the strength and frequency of their presumed deviant sexual interests. In the last several years, however, this heightened focus on the role of deviant sexual arousal has shifted considerably; in part because it is now recognized that most adolescents who commit a sexual offense do not display primarily deviant sexual interests. For example, in their analysis of archival penile plethysmograph (PPG) data, Seto, Lalumière, and Blanchard (2000) reported that 25% (or 10 out of 40) of a small sample of adolescents who had offended sexually demonstrated maximal sexual interest in prepubescent children. With an overlapping yet augmented sample, Seto, Murphy, Page, and Ennis (2003) found that 30% (76 out of 253) of adolescents who had offended sexually responded equally or more to child stimuli during PPG assessments. In a study of 136 adolescents who had committed a sexual offense, Worling (2004) found that 36% of the adolescents were rated by clinicians as having sexual interests in prepubescent children and/or sexual violence. Although deviant sexual interests likely play a role in the etiology and/or maintenance of adolescent sexual offending for *some* adolescents, there are likely other factors to consider such as intimacy deficits, antisociality, attitudes supportive of sexual offending, and opportunity, for example.

Measurement of Sexual Arousal

Given the assumption that individuals will be reluctant to reveal their sexual thoughts and feelings during an interview, some believe that it is necessary to utilize a physiological measure of sexual interest. The most popular technique for measuring sexual arousal with adult males is the PPG. Although it has been argued by some that the PPG can provide valuable information for adult male clients who have offended sexually (Lalumière & Harris, 1998; Seto, 2001), there are some concerns regarding the reliability and validity of PPG data collected from adult males (Konopasky & Konopasky, 2000; Marshall & Fernandez, 2000). There are also several ethical and empirical concerns regarding the use of the PPG with adolescents (Becker & Harris, 2004; Hunter & Lexier, 1998; Worling, 1998). For example, adolescence is a period of emerging sexual development and teenagers are continuously developing and refining their sexual scripts, identities, and preferences. The potential for iatrogenic harm from exposing adolescents to visual and/or auditory depictions of deviant sexual activities is considerable, yet this has never been examined. Second, there is little compelling evidence regarding the reliability or validity of a PPG procedure with adolescents. Indeed, in studies with adolescents, PPG data are significantly influenced by such variables as the adolescent's age (Kaemingk, Koselka, Becker, & Kaplan, 1995) and history of physical and sexual abuse (Becker, Hunter, Stein, & Kaplan, 1989; Becker, Kaplan, & Tenke, 1992). Furthermore, although PPG data are predictive of subsequent sexual offending for adults (Hanson & Bussière, 1998; Hanson & Morton-Bourgon, 2004), the available research with adolescents indicates that there is no significant relationship between phallometrically-measured sexual deviance and sexual assault recidivism (Gretton, McBride, Hare, O'Shaughnessy, & Kumka, 2001; Gretton et al., 2005). Becker et al. (1992) also found that a majority of adolescents who denied their sexual offenses provided invalid deviant-arousal data. In a chart-review study by Seto et al. (2000), it was found that PPG data were only moderately discriminative

for adolescent males who sexually assaulted male victims and that adolescents who offended against female children could not be differentiated from a nonoffending population. Hunter, Becker, and Goodwin (1994) similarly found that only those adolescents with male victims demonstrated significant deviant arousal using the PPG. In their review of the research, Becker and Harris (2004) concluded that the PPG may be useful *only* for the following adolescent clients: (i) older teenage males, (ii) individuals who acknowledge their sexual assaults, and (iii) those who offend sexually against males.

Given the scientific and ethical concerns regarding phallometric assessment, a number of researchers and clinicians have begun to use an alternative physiological measure: unobtrusively measured viewing time (VT). With this procedure, clients are asked to rate the sexual attractiveness of photographs of a variety of models while the response time to provide the ratings is unobtrusively recorded. The assumption underlying this technique is that people will look longer at stimuli that they find sexually attractive relative to stimuli that they find sexually unattractive. In several VT assessment systems that are commercially available, the models in the photographs are clothed and are not displayed in sexual poses. For many clinicians, therefore, the VT assessment addresses some of the ethical concerns raised by the PPG.

In studies with nonoffending adults, it has been demonstrated that VT is significantly correlated with self-reported ratings of sexual arousal (e.g., Harris, Rice, Quinsey, & Chaplin, 1996; Lang, Searles, Lauerman, & Adesso, 1980; Quinsey, Ketsetzis, Earls, & Karamanoukian, 1996) and with sexual arousal as measured by the PPG (Harris et al., 1996; Quinsey et al., 1996). There have also been several investigations of the efficacy of VT with adults who have offended sexually. For example, Harris et al. (1996) found that men who sexually assaulted children viewed slides of children longer than they viewed slides of adults. In several studies conducted with the Abel Assessment for Sexual Interest™ (AASI), a VT assessment methodology using photographs of clothed models, Abel and his colleagues have reported encouraging data with respect to the internal consistency and discriminant validity of the AASI with adult males who acknowledged sexual assaults against children (Abel, Jordan, Hand, Holland, & Phipps, 2001; Abel, Lawry, Karlstrom, Osborn, & Gillespie, 1994; Abel, Huffman, Warberg, & Holland, 1998). In one of the few studies of the utility of the AASI conducted outside of the developer's laboratory, Letourneau (2002) found that PPG and AASI responses from 57 adult males were significantly correlated for most stimulus categories, and that both assessment approaches significantly differentiated those participants with male child victims from those with victims from other groups. Furthermore, the AASI VT procedure could identify those men with adolescent female victims. More recently, Gray and Plaud (2005) compared the AASI and the PPG using data collected from 39 men who had offended sexually against a child less than 11 years of age. They found that both procedures significantly identified sexual interest in children.

There have been few published studies of the utility of VT with adolescents. In one paper, Smith and Fisher (1999) used Abel's VT system (AASI) with 81 adolescent males. They concluded that this particular approach yielded little convincing evidence with respect to reliability or validity. However, it is important to note that Abel (2000) later disputed many of the results and interpretations presented by the investigators. More recently, Abel et al. (2004) evaluated the AASI with data collected from 1,704 males aged 11 to 17. The authors reported that VT for images of children was moderately correlated to the number of child victims ($r = .18$) and the number of acts of child sexual offending ($r = .23$). It was also noted that VT for child stimuli could moderately differentiate those adolescents who offended sexually against children from those who offended against peers or adults (AUC = .64).

Despite the prevailing assumption that self-reported sexual interests from individuals who commit sexual offenses will be necessarily biased and prejudicial, there is growing evidence of the utility of self-reported data. For example, Laws, Hanson, Osborn, and Greenbaum (2000) found that self-reported sexual interests showed more classification accuracy than PPG data with respect to victim gender in a sample of men who offended against children. Laws et al. also found that the self-report methodology used (card sort) had excellent levels of internal consistency. In a similar study, Day, Miner, Sturgeon, and Murphy (1989) found that self-report data from a structured questionnaire regarding sexual thoughts, feelings, and behaviors could accurately classify men according to the gender of their child victims.

Turning to research with adolescents, Seto et al. (2000) found that most participants who acknowledged a sexual interest in children were classified as "pedophilic" using the PPG. Furthermore, Seto et al. found that those participants who offended sexually against children but who denied sexual interest in children scored significantly lower on the pedophilic index. Daleiden, Kaufman, Hilliker, and O'Neil (1998) found that adolescents who offended sexually disclosed significantly more deviant sexual behaviours relative to both nonsexual offenders and a nonoffending group. Results such as these suggest that individuals may be quite open regarding deviant sexual thoughts, interests, and/or behaviors.

Method

Participants

After obtaining informed consent (and parental consent when required), data were collected from 78 males aged 12–18 ($M = 15.09$; $SD = 1.62$). Participants were assessed at a residential treatment facility in Minnesota ($n = 44$) or at one of three community-based treatment centers in the Greater Toronto (Ontario, Canada) area ($n = 34$). Females were not specifically excluded; rather the facility in Minnesota provides services only to males, and the 34 consecutive Canadian referrals who chose to participate were male.

All participants for this study acknowledged a contact sexual offense. During the course of this investigation, none of the adolescents who completely denied their sexual offenses and who were approached to participate ($n = 5$) volunteered for this study. Sixty-seven percent (52/78) of the adolescents committed a sexual offense against at least one child (defined as under the age of 12 and 4 or more years younger than that adolescent at the time of the offense); the remainder (33% or 26/78) of the adolescents committed sexual assaults against peers or adults exclusively (5 of these 26 offended against male peers). With respect to the participants' ethnic origin, 81% (63/78) were Caucasian, 12% were African American/Canadian, 5% (4/78) were native American/Canadian, and 2% (2/78) were Hispanic. Information regarding intellectual functioning was available for 52 of the 78 adolescents: 19% (10/52) of these adolescents had IQ's less than 80, 6% (3/52) of the adolescents had IQ's above 120, and 75% (39/52) had an IQ's within the average range (i.e., IQ between 80 and 120). Assessors reported that 50% (39/78) of the participants disclosed a childhood sexual victimization history and 42% (33/78) of the adolescents disclosed physical abuse within their families.

Self-Report Sexual Arousal Graphs

To collect self-reported sexual arousal data, adolescents were asked to complete 2 graphs (1 graph for each gender) on which they rated their sexual arousal for 8 age-based categories:

0–3 years, 4–6 years, 7–9 years, 10–12 years, 13–15 years, 16–18 years, 18–24 years, and over 24 years (see Appendix A for self-report arousal graphs and instructions). Adolescents were first asked to rate their level of sexual arousal to the various age groups if there was no physical force or violence involved in the sexual interaction; they were then asked to use a different color of pen and indicate how their sexual arousal would change for each age group when the thought of forced sexual contact was introduced. All ratings were made on a scale from 0 (low) to 10 (high). For all participants, the graph to address sexual arousal to males was completed first, followed by the graph to capture sexual arousal to females. The graphing procedure was completed with the adolescents by a psychologist or social worker during the completion of a comprehensive assessment. There were 7 different clinicians (1 in Minnesota and 6 in Ontario) at the various agencies who assisted the adolescents to complete the graphs.

Affinity Assessment of Sexual Interest

The Affinity procedure is explained in detail elsewhere (Glasgow, Osborne, & Croxen, 2003), and assessors followed the assessment protocol outlined in the Affinity 1.0 manual (Glasgow, 2001). Briefly, the assessment involves the computerized presentation of photographs of 28 males and 28 females in 4 age categories: toddlers, preadolescents, adolescents, and adults. All of the photographs are of clothed individuals, and none of the models are depicted in sexual poses. The 56 images are presented in a fixed, random order, and participants are asked to rate the sexual attractiveness of each image using a mouse pointer. The on-screen self-report rating scale accompanying each photograph includes the anchors "Very unattractive," "Neutral," and "Very attractive," and there are 19 different, unnumbered gradients on the scale. The Affinity computer program registers a self-report rating score for each photograph ranging from 0 to 18, and it records the time taken (in seconds; accurate to ± 0.02 s) to provide each attractiveness rating. The later measure is referred to as on-task latency (OTL). Immediately prior to the debriefing procedure where the nature of the Affinity program was explained, participants were asked to rate—on a scale from 1 (*Not at all upsetting*) to 10 (*Very upsetting*)—how upsetting they found the experience of rating the Affinity photographs. They were also asked to indicate how enjoyable the experience was using a scale from 1 (*Not at all enjoyable*) to 10 (*Very enjoyable*), and participants' comments regarding the assessment process were recorded by the assessor.

Average Sexual Arousal Scores

As a result of the fact that the Affinity program uses four age groups (toddler, preadolescent, adolescent, and adult) for each gender, the eight age groups examined using the self-report graphing procedure were collapsed to approximate the same four Affinity groups to facilitate raw-score comparisons. Specifically, self-report sexual arousal graph data were collapsed as follows: ages 0–3 and 4–6 (toddler); ages 7–9 and 10–12 (preadolescent); 13–15 and 16–18 (adolescent); 18–24 and 24 and over (adult).

To facilitate visual comparisons between the three assessment techniques, raw scores were converted to z scores. The average standardized raw scores for the three assessment approaches are displayed in Fig. 1. The raw data distributions are remarkably similar across approaches, despite the differences in measurement scales and assessment methodology. Specifically, average self-reported sexual arousal ratings and VT scores for female adults and adolescents are significantly higher than the arousal data for all remaining stimulus

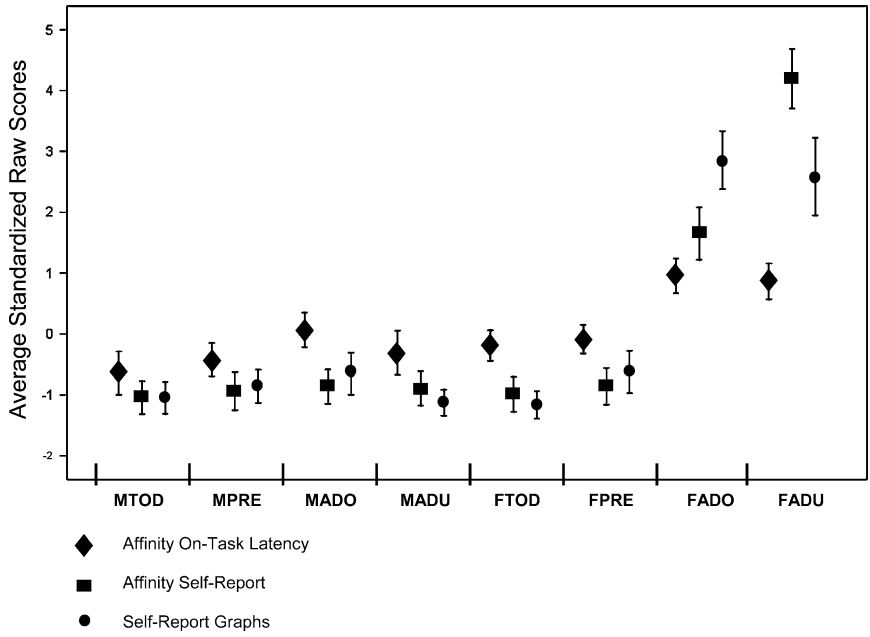


Fig. 1 Average standardized raw scores for Affinity On-Task Latency (OTL; $n = 78$), Affinity Self-Report ($n = 78$), and Self-Report Sexual Arousal Graph ($n = 72$) procedures across stimulus groups. Bars represent 95% confidence intervals. MTOD: Male Toddler; MPRE: Male Preadolescent; MADO: Male Adolescent; MADU: Male Adult; FTOD: Female Toddler; FPRE: Female Preadolescent; FADO: Female Adolescent; FADU: Female Adult

categories. Furthermore, there was significant overlap in the arousal data provided for males of any age and for female toddlers and preadolescents.

Internal Consistency

Internal consistency estimates (Cronbach's α) for the 8 Affinity stimulus categories are displayed in Table 1. With the exception of the α for OTL measure for the 7 photographs of female adolescents, all internal consistency estimates were above .70 for the OTL values across the age/gender categories. Note, however, that higher internal consistency estimates were found for all Affinity stimulus categories based on self-report ratings.

Internal consistency estimates (Cronbach's α) for the self-report sexual arousal graphs were .83 for females (collapsed across the four age groups) and .89 for males (collapsed across the four age groups). For children 12 and under (for both males and females), the internal consistency estimate was .92; for the 4 age groups of both males and females over the age of 12, internal consistency was .77.

Validity

The assumption underlying VT is that, when asked to rate the sexual attractiveness of a particular model, individuals look longer at photographs that they find sexually arousing relative to the time that they spend looking at photographs that they find not sexually

Table 1 Internal consistency estimates for Affinity on-task latency and Affinity self-report ratings

Stimulus category	Affinity on-task latency (measured in seconds) Cronbach's α	Affinity self-report ratings (scale from 0–18) Cronbach's α
Female toddlers	.82	.96
Female preadolescents	.79	.95
Female adolescents	.62	.87
Female adults	.72	.94
Male toddlers	.73	.97
Male preadolescents	.82	.97
Male adolescents	.77	.94
Male adults	.77	.94

arousing. The correlations pertaining to this issue are displayed in Table 2, and it can be seen that for 7 of the 8 stimulus categories, the time spent viewing the slides (Affinity OTL) was significantly correlated with self-report ratings for the same slides (Affinity Self-Report). The average correlation between Affinity OTL and Affinity self-report ratings across all slide categories was $r = .31$, $p < .01$. There was, however, a significant, *negative* correlation between OTL and Affinity Self-Report for the slides of the adult females. In other words, there was a marked tendency for those adolescents who rated the photographs of adult females as sexually attractive to provide their ratings more quickly than those adolescents who found the photographs less sexually attractive. This, of course, is counter to the assumption that forms the premise of the VT approach.

Although the Affinity self-report is based on photographs of various models, and the self-report sexual arousal graphs simply require a rating in response to a numeric age group, the correlations between these two measures of sexual arousal were significant for all 8 stimulus categories. The correlations between OTL and the self-report graphs were also significant for 6 of the 8 stimulus categories.

The validity of the two Affinity procedures (OTL and self-report) and the self-report graphing procedures was examined by comparing subgroups of adolescents formed on the basis of the age and gender of their victims. First, an Affinity OTL deviance index was calculated for each participant by dividing the highest mean OTL (in seconds) for male or female toddlers or preschoolers by the highest mean for male or female adolescents or adults.[1] An Affinity Self-Report deviance index was calculated using the same mathematical procedure: the highest mean Affinity self-report rating for male or female toddlers or preschoolers was divided by the highest mean self-report rating for male or female adolescents or adults. Finally, a deviance index was calculated in the same fashion for the self-report graphing procedure by dividing the highest rating for a male or female child aged 0 to 12 by the highest rating for a male or female aged 13 and over.

To examine the discriminative validity of the three assessment procedures, the area under the Receiver Operating Characteristic (ROC) curve (AUC) was calculated for several different between-group comparisons (see Table 3). Both of the self-report procedures could significantly identify those adolescents with single child victims, multiple child victims, or male child victims (whether defined as ever a male child or exclusively male child victims),

[1] Deviance indices were also calculated for each of the three assessment techniques using a difference score: i.e., the highest mean response to male or female toddlers or preadolescents minus the highest mean response to male or female adolescents or adults. However, this procedure resulted in very poor discriminative power for most analyses for all three assessment techniques.

Table 2 Correlations between Affinity on-task latency, Affinity self-report, and self-report sexual arousal graph procedures

Stimulus group	Affinity self-report	Self-report sexual Arousal graphs
Male Toddler		
Affinity OTL	.38**	.35**
Affinity self-report		.64**
Male preadolescent		
Affinity OTL	.67**	.55**
Affinity self-report		.79**
Male Adolescent		
Affinity OTL	.61**	.46**
Affinity self-report		.63**
Male adult		
Affinity OTL	.52*	.21*
Affinity self-report		.50**
Female toddler		
Affinity OTL	.49**	.44**
Affinity self-report		.80**
Female preadolescent		
Affinity OTL	.48**	.38**
Affinity self-report		.73**
Female Adolescent		
Affinity OTL	.24*	.16
Affinity self-report		.34**
Female Adult		
Affinity OTL	− .26*	− .17
Affinity self-report		.48**

Note. OTL: On-task latency. $n = 78$ for all correlations between Affinity self-report and Affinity OTL. $n = 72$ for all correlations with the Self-report sexual arousal graphs.

*$p < .05$; ** $p < .01$.

and there were no significant differences between the two self-report methodologies, all Z_Δ's $\leq$ .03, all p's $\leq$ .05. The Affinity OTL deviance index was only able to discriminate those adolescents with a male child victim from those adolescents with victims from all other age/gender groups, and neither self-report measure was superior to the Affinity OTL for this discrimination, all Z_Δ's $\leq$ .03, all p's $>$ .05. Interestingly, none of the deviance indices examined correctly identified those adolescents with female child victims: whether this was defined as if ever a female child victim or exclusively female child victims. It should also be pointed out that of the 30 adolescents who had 2 or more child victims, 23/30 (or 77%) had at least one male child victim. Only 7/30, or 23%, had 2 or more child victims that were exclusively female.

Given that there are few published data using the Affinity assessment system, one-way ANOVAs were also computed for the 3 deviance measures for each of the comparisons described above. The means and standard deviations for these comparisons are presented in Table 4 to facilitate the calculation of effect sizes and comparisons with future investigations. The pattern of group differences is similar to that found with the ROC analyses. Also note in Table 4 that there were no significant differences on any of the deviance measures as a function of sexual or physical victimization history or location of data collection.

Table 3 Area under the ROC curve (AUC) classification data for three measures of deviant sexual interest

Sexual assault victim(s)	Affinity on-task latency deviance index AUC ($n = 78$)	95% CI	Affinity self-report deviance index AUC ($n = 78$)	95% CI	Self-report sexual arousal graph deviance index AUC ($n = 72$)	95% CI
Ever a child victim	.61	.47–.75	.67*	.54–.79	.66*	.53–.79
2 or more child victims	.60	.47–.73	.73**	.61–.86	.70**	.59–.82
Ever a male child victim	.69**	.57–.81	.72**	.60–.84	.72**	.60–.84
Only male child victim(s)	.73**	.60–.86	.74**	.60–.89	.76**	.64–.88
Ever a female child victim	.42	.30–.55	.48	.35–.61	.45	.33–.58
Only female child victim(s)	.43	.29–.56	.42	.28–.56	.41	.27–.56

Note. Affinity On-Task Latency Deviance Index = (highest mean on-task latency for male or female toddlers or preadolescents)/(highest mean on-task latency for male or female adolescents or adults). Affinity Self-Report Deviance Index = (highest mean self-reported ratings for male or female toddlers or preadolescent)/(highest mean self-reported ratings for male or female adolescents or adults). Self-Report Sexual Arousal Graph Deviance Index = (highest rating for male and female children aged 0–3, 4–6, 7–9, and 10–12)/(highest rating for males and females aged 13–15, 16–18, 18–24, and 24 and over).

*$p < .05$; **$p < .01$.

Table 4 Mean (and *SD*) deviance indexes for between-group comparisons

Comparison	Affinity on-task latency deviance index ($n = 78$)	F	Affinity self-report deviance index ($n = 78$)	F	Self-report sexual arousal graph deviance index ($n = 72$)	F
Ever a child victim						
Yes	.72 (.20)	1.6	.17 (.27)	2.2	.29 (.29)	1.79
No	.65 (.24)		.08 (.24)		.19 (.32)	
2 or more child victims						
Yes	.74 (.20)	1.6	.24 (.31)	7.4 **	.34 (.25)	4.1*
No	.67 (.23)		.08 (.20)		.19 (.31)	
Ever a male child victim						
Yes	.77 (.20)	7.1 **	.24 (.31)	7.4 **	.37 (.28)	7.1 **
No	.65 (.21)		.08 (.20)		.18 (.29)	
Only male child victim(s)						
Yes	.84 (.20)	10.9 **	.32 (.37)	11.7 **	.39 (.29)	4.2*
No	.66 (.21)		.09 (.19)		.20 (.29)	
Ever a female child victim						
Yes	.66 (.18)	1.9	.10 (.16)	1.6	.24 (.27)	0.2
No	.73 (.24)		.18 (.31)		.27 (.32)	
Only female child victim(s)						
Yes	.64 (.19)	0.1	.08 (.13)	0.2	.17 (.26)	0.2
No	.72 (.22)		.17 (.29)		.28 (.31)	
Participant a victim of sexual abuse						
Yes	.68 (.22)	0.4	.20 (.03)	0.9	.24 (.29)	0.7
No	.72 (.21)		.31 (.05)		.27 (.30)	
Participant a victim of physical abuse						
Yes	.67 (.23)	0.7	.13 (.22)	.07	.20 (.25)	1.7
No	.71 (.21)		.15 (.28)		.29 (.32)	
Location of participant						
Minnesota	.70 (.22)	0.0	.23 (.03)	0.1	.26 (.31)	0.2
Greater Toronto area	.70 (.22)		.33 (.07)		.23 (.28)	

Note. Affinity On-Task Latency Deviance Index = (highest mean on-task latency for male or female toddlers or preadolescents) / (highest mean on-task latency for male or female adolescents or adults). Affinity Self-Report Deviance Index = (highest mean self-reported ratings for male or female toddlers or preadolescent)/(highest mean self-reported ratings for male or female adolescents or adults). Self-Report Sexual Arousal Graph Deviance Index = (highest rating for male and female children aged 0–3, 4–6, 7–9, and 10–12)/(highest rating for males and females aged 13–15, 16–18, 18–24, and 24 and over).

*$p < .05$; **$p < .01$.

Table 5 Correlations between deviance indices and number of known child victims

Variable	Affinity on-task latency deviance index ($n = 78$)	Affinity self-report deviance index ($n = 78$)	Self-report sexual arousal graphs deviance index ($n = 72$)
Number of child victims	.07	.23*	.14
Number of male child victims	.33**	.50**	.34**
Number of female child victims	−.13	−.04	−.05

Note. Affinity On-Task Latency Deviance Index = (highest mean on-task latency for male or female toddlers or preadolescents)/(highest mean on-task latency for male or female adolescents or adults). Affinity Self-Report Deviance Index = (highest mean self-reported ratings for male or female toddlers or preadolescent)/(highest mean self-reported ratings for male or female adolescents or adults). Self-Report Sexual Arousal Graph Deviance Index = (highest rating for male and female children aged 0–3, 4–6, 7–9, and 10–12)/(highest rating for males and females aged 13–15, 16–18, 18–24, and 24 and over).

*$p < .05$; **$p < .01$.

It should also be pointed out that there was no significant correlation between participant age and the deviance index calculated for Affinity OTL ($r = .02$), Affinity Self-Report ($r = .09$), or the self-report sexual arousal graphs ($r = -1.1$), all p's $> .05$. Correlations between the three deviance measures and the number of known child victims are displayed in Table 5.

Combinatory deviance scores were also computed to determine whether the results from multiple assessment methods could enhance discriminatory power. Specifically, the following combined deviance scores were examined: Affinity OTL + Affinity self-report; Affinity OTL + self-report arousal graph; Affinity self-report + self-report arousal graph; and Affinity OTL + Affinity self-report + self-report arousal graph. For the six between-groups comparisons outlined in Table 3, there was no significant increase in predictive efficiency for any of the combinatory deviance scores, all Z_{Δ}'s $\leq .06$, all p's $> .05$. The highest AUC value obtained for any comparison was 0.79 (95% CI; .66–.91), and this represented the ability to detect multiple child victims using the combination of Affinity self-report plus the self-report sexual arousal graphing procedure.

Given the relative novelty of the self-report sexual arousal graphs, a final examination of the validity of this technique involved an analysis of the frequency with which adolescents *ever* rated their sexual arousal over the "0" line for male or female children aged 0 to 12 years. The data related to this analysis are presented in Table 6, and it can be seen that if an adolescent ever provided a rating over "0" for children aged 0 through 12 years, this was significantly related to the choice of a single child victim, multiple child victims, and male child victims. As with all previous analyses, however, this particular procedure could not discriminate those adolescents with female child victims.

Finally, participants were asked to rate how upset they were after completing the Affinity rating task. On the scale from 1 (*Not at all upsetting*) to 10 (*Very upsetting*), the mean rating was 2.38 ($SD = 1.88$). Only 6 adolescents provided a rating of 5 or higher, and 66% of the participants provided a rating of 1 or 2. Anecdotally, those adolescents who rated the Affinity procedure as somewhat upsetting most often commented that there were too many photographs of young children or that there were too many photographs of males. Participants were also asked to rate how enjoyable they found the Affinity computerized assessment using a scale from 1 (*Not at all enjoyable*) to 10 (*Very enjoyable*). The mean

Table 6 Frequency of self-report sexual arousal graph ratings over "0" for children aged 0 to 12

	Ever more than "0" sexual arousal		
	No	Yes	Total
Ever a child victim			
No	16	9	25
Yes	15	32	47
Total	31	41	72
$\chi^2 = 6.9^*$			
Ever 2 or more child victims			
No	26	19	45
Yes	5	22	27
Total	31	41	72
$\chi^2 = 10.6^{**}$			
Ever a male child victim			
No	26	18	44
Yes	5	23	28
Total	31	41	72
$\chi^2 = 11.87^{**}$			
Only male child victim(s)			
No	28	27	55
Yes	3	14	17
Total	31	41	72
$\chi^2 = 5.86^*$			
Ever a female child victim			
No	19	23	42
Yes	12	18	30
Total	31	41	72
$\chi^2 = 0.20$			
Only female child victim(s)			
No	21	32	53
Yes	10	9	19
Total	31	41	72
$\chi^2 = 0.97$			

Note. Ratings were provided on a scale from 0 (low) to 10 (high).

$^*p < .05$; $^{**}p < .01$.

rating for this variable was 3.70 ($SD = 2.10$). Those adolescents who found the procedure enjoyable most often commented that it was "interesting"; those who rated the Affinity assessment as *Not at all enjoyable* most often mentioned that it was "boring."

Discussion

Adolescent sexual interest was assessed using three different assessment procedures: the Affinity (version 1.0) VT procedure, the Affinity self-report procedure, and a self-report sexual arousal graphing procedure. Overall, the internal consistency estimates for all three measures were acceptable for most age and gender groups, and all three assessment approaches significantly differentiated those adolescents with male child victims from those adolescents who never offended sexually against a male child. On the other hand, none of the

assessment techniques could successfully differentiate adolescents with female child victims from those who offended sexually against other groups.

The finding that those adolescent males with male child victims demonstrated the highest interest in prepubescent children—either through VT or self-report—is consistent with results from several studies in which the PPG was used (Becker et al., 1989; Hunter et al., 1994; Murphy, DiLillo, Haynes, & Steere, 2001; Seto et al., 2000). As has been noted by many of these researchers, it is quite likely, therefore, that many adolescents who commit a sexual offense against a male child are motivated, at least in part, by a sexual interest in prepubescent children. Of course, not all adolescents who target male children demonstrate deviant sexual arousal, and other factors, such as opportunity or antisocial attitudes, for example, need to be considered in the etiology and/or maintenance of their sexual offenses.

Similarly, the fact that many participants who offended sexually against female children did not demonstrate or report significant sexual interest in prepubescent children is also consistent with PPG data collected from adolescents (e.g., Seto et al., 2000) and with both PPG and VT data collected from adults (e.g., Letourneau, 2002). As has been noted by others, this may be a result of the fact that the various assessment procedures are not yet able to discern such a sexual interest. Alternatively, given the fact that this finding has been observed with both adults and adolescents—using PPG, VT, and self-report—it is possible that there are many males who commit sexual offenses against prepubescent girls for reasons other than deviant sexual interest. As noted with respect to some of those who offend against boys, it could be that factors such as opportunity or abuse-supportive attitudes, for example, are critical in the formation and/or continuation of offending behaviors.

The AUC reported by Abel et al. (2004) for the AASI to differentiate those adolescents who ever assaulted a child from those who sexually assaulted a peer or adult was .64. The AUC for the Affinity VT in the present study (AUC = .61) was remarkably similar for the same comparison. As noted clearly by Abel et al. (2004), this suggests that VT alone should not be used to classify adolescents who have offended sexually with respect to victim age. Of course, despite the fact that both of the self-report procedures significantly differentiated adolescents with versus without child victims, no one assessment methodology produced an AUC over .74 regardless of the discrimination examined, and no combinatory deviance score produced an AUC over .79. This indicates that there is certainly much room for improvement, and that neither self-report procedure should be used on its own, or in combination, to determine whether or not an adolescent has committed a sexual offense against a person from a specific age/gender category. Of course, it is also possible that researchers will never find that they can get better discrimination between subgroups on the basis of victim gender and/or age because of the fact that sexual interests and attitudes are still being formed during adolescence. Although victim gender and/or age may be proxy markers of sexual interest for some adults who offend sexually, particularly those who offend against boys, it is likely that the choice of victim is less often related to fixed sexual interests for adolescents.

Despite the encouraging results reported regarding both self-report procedures, it should be stressed that there were no adolescents who completely denied their sexual crimes included in this study. As such, these results should not be generalized to that subgroup. It would be interesting to investigate the utility of both VT and self-report with those who deny their sexual offenses as Becker et al. (1992) found that the PPG was not useful with this group as most of these adolescents were classified as "nonresponders." There were also no female participants included in this study. Given the relative ease with which both Affinity and the self-report graphing procedures can be carried out with both male and

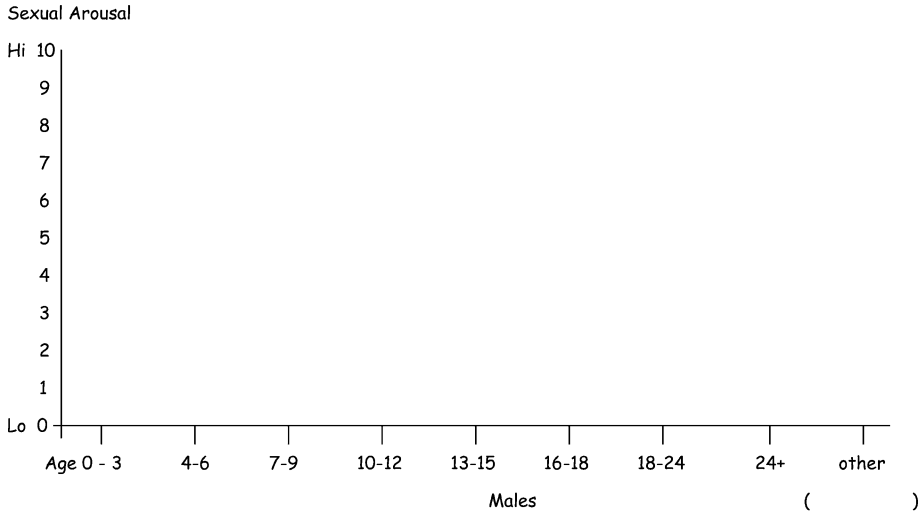
female adolescents, it would be beneficial to learn about the psychometric properties of these assessment techniques with female participants. It would also have been ideal to have included a group of adolescents without a history of sexual offending. This would have permitted analyses regarding normative responses for both VT and self-report, and there may have been greater discrimination accuracy for specific subgroups if nonoffending youth were included. Finally, it should also be pointed out that there were only 78 participants in this investigation; therefore, the probability of a Type 1 error should be considered given the number of statistical tests herein.

Although VT was significantly correlated with ratings of sexual interest for the same photographs within most of the stimulus categories, there was a significant negative correlation between VT and sexual attractiveness ratings for photographs of adult females. It is difficult to comment on the uniqueness of this finding, however, as few researchers have published correlations between self-reported ratings and VT scores for separate stimulus categories. Rather, it is more common for researchers to report an average correlation across the various age/gender categories ($r = .31$ in the present study). One can quickly see in Fig. 1 that, on average, adolescents viewed photographs of adolescent and adult females for approximately the same amount of time, but that the sexual attractiveness ratings provided for the photographs of the adult females were significantly higher than those ratings given to the photographs of adolescent females. Given that the relationship between VT and self-reported attractiveness ratings were in the expected direction for most stimulus categories, it may be that this result is spurious and sample specific. On the other hand, there may be something about the seven photographs of adult women used in the Affinity program that contributed to this unexpected result. Alternatively, it is possible that some participants provided exaggerated ratings to the slides depicting adult females in an effort to appear more socially desirable. Finally, it is possible that there is an inverted-U-shaped distribution of VT responses such that respondents provide the fastest ratings to categories that they find both highly attractive and highly unattractive. Given that the majority of adolescents in this study found only 2 of the 8 categories of photographs as highly sexually attractive (adolescent and adult females), the fast VT responses to the 6 remaining categories would account for an overall positive correlation between low VT and low attractiveness ratings. In other words, participants provided low average attractiveness ratings after low average viewing times for 75% (42/56) of the slides. It would be ideal for researchers to examine the possibility of such an inverted-U-shaped distribution in future research with VT technology; particularly given the assumption underlying the technique that there is a linear relationship between viewing times and attractiveness ratings for *all* age/gender categories.

Most of the adolescents found that participating in the Affinity VT assessment was not very upsetting, and many even found the process to be at least somewhat enjoyable. Given the potential intrusiveness of the PPG procedure, it is interesting that researchers have never collected information from adolescents regarding their experience of providing PPG data. Given the current scientific limitations of the PPG with adolescents, and the ethical concerns involved in conducting such an assessment, it is encouraging that at least one VT approach (Affinity 1.0) can produce some useful information regarding sexual interests without significant negative impact. Furthermore, it is important to point out that age at time of assessment, childhood sexual victimization history, and physical abuse history were not significantly related to deviance indices calculated for Affinity VT, Affinity Self-Report, or the self-report graphing procedure. Recall that some researchers have found that PPG data are correlated with these factors.

The results of the present study also lend support to the growing evidence that structured, self-report methodologies can provide the assessor with valuable information regarding an adolescent's sexual interests. For example, adolescents with a history of sexual offenses who rated their sexual arousal to children 12 years of age and under over the "0" mark on the self-report arousal graphs were significantly more likely to have offended against a single child victim, multiple child victims, or male child victims. Of course, regardless of the particular assessment methodology chosen, it is always prudent to supplement any sexual-interest data with information from official documentation and with reports from parents and other adults who may be familiar with the adolescent. It is also essential to be mindful of the fact that sexual interests are not necessarily crystallized for many adolescents and, furthermore, that victim choice may not necessarily be a proxy for sexual preferences. Even when deviant sexual interests are identified, however, there is very little research regarding how best to help adolescents in this regard. Although covert sensitization is by far the most popular behavioral treatment for deviant-arousal reduction in programs in the United States for adolescents who have offended sexually (McGrath, Cumming, & Burchard, 2003), there are no published data regarding the efficacy of this approach for adolescents. In addition to the continued refinement of measurement strategies, therefore, it is critical that we examine the impact of treatment and management techniques for those adolescents who evidence deviant sexual interests.

Appendix A: Self-Report Sexual Arousal Graphs and Instructions



Self-Report Sexual Arousal Graphs: Instructions

While sitting at a table at 90 degrees from the adolescent, place the first graph in front of the adolescent—start with the graph (male or female) that is the same gender as the adolescent. That is, for male adolescents start with the "Males" graph.

> This is a special graph to help us look at the kinds of things that you would find sexually arousing; that is, things that would turn you on in a sexual way. Everyone is different regarding what kinds of things turn them on and turn them off, and it is important to go through this slowly. It is also really important that you are as honest as possible.

(point to the first graph) "You see that this graph has a scale from 0 to 10. Zero means that you would not be sexually aroused at all—that you would not find a person in that category sexually attractive at all. You might even think that it would be gross or disgusting. A ten means that you would be really sexually aroused or turned on sexually. A five would mean that you would be somewhat turned on sexually, and so on. Down on the bottom are males [females] of different age groups. Now the first group of ratings we are going to make are ratings of how turned on you would be doing sexual things with males [females] without the use of force—i.e., not pushing, or hitting, and no other violence; just doing sexual things and the person would go along with you."

(hand them a multi-color pen—you could also just have available several different colors of pens/pencils) "Take this pen and choose a color and color in this box here that says 'No Force' because this will be the legend for our graph." (*interviewee takes pen* and colors in the "No Force" legend with the colored pen). "Now, with that same color, show me by putting a mark on this graph how sexually aroused you would be doing sexual things with a male aged 0-3, so that would be boys who are newborn, and 1, 2, and 3 years old. Remember, everyone is different regarding what kinds of people turn them on. Now, what about your sexual arousal to males aged 4-6?" Continue through the age groups slowly. When it comes to the "other" category, this is when one can ask about arousal to family members or other specific sexual targets that you want to question (e.g., animals).

"Now, choose a different color and color in the 'Force' box. Now we are going to rate how sexually turned on or aroused you would be if you used force to do sexual things with people. By 'force,' I mean holding someone down or hitting them and making them do sexual things when they don't want to. Everyone is different, and some people are more sexually turned on when they think of using force and some people are more turned off. For others it doesn't really change how turned on they would be. How does the use of force change how sexually turned on you would be doing sexual things with a male aged 0 to 3?" Then proceed through other age groups.

This whole procedure is then repeated with the graph for the second gender.

Acknowledgements This research was supported, in part, by a grant from Youth Justice Services, Central Regional Office, Ontario Ministry of Children and Youth Services. The views expressed herein are mine and do not necessarily reflect those of the Thistletown Regional Centre for Children and Adolescents or the Ontario Ministry of Children and Youth Services. I am indebted to the adolescents who took part in this investigation and to Sarah Herrick (formerly of Mille Lacs Academy/Nexus) for her role in assessment and data collection. I am also very grateful for the tremendous support of this research by Kathe Dellacecca, Director of Mille Lacs Academy/Nexus and by Barbara Rodgers, Director of the SAFE-T Program. Special thanks also to Sabrina Ramdeholl for her valuable comments on a previous draft of this manuscript.

References

Abel, G. G. (2000). The importance of meeting research standards: A reply to Fisher and Smith's articles on the Abel Assessment for Sexual Interest™. *Sexual Abuse: A Journal of Research and Treatment*, *12,* 155–161.

Abel, G. G., Huffman, J., Warberg, B., & Holland, C. L. (1998). Visual reaction time and plethysmography as measures of sexual interest in child molesters. *Sexual Abuse: A Journal of Research and Treatment*, *10,* 81–95.

Abel, G. G., Jordan, A., Hand, C. G., Holland, L. A., and Phipps, A. (2001). Classification models of child molesters utilizing the Abel Assessment for Sexual Interest(tm). *Child Abuse & Neglect*, *25,* 703–718.

Abel, G. G., Jordan, A., Rouleau, J. L., Emerick, R., Barboza-Whitehead, S., & Osborn, C. (2004). Use of visual reaction time to assess male adolescents who molest children. *Sexual Abuse: A Journal of Research and Treatment*, *16,* 255–265.

Abel, G. G., Lawry, S. S., Karlstrom, E. M., Osborn, C. A., & Gillespie, C. F. (1994). Screening tests for pedophilia. *Criminal Justice and Behavior*, *21,* 115–131.

Becker, J. V., & Harris, C. (2004). The psychophysiological assessment of juvenile offenders. In G. O'Reilly, W. L. Marhsall, A. Carr, & R. C. Beckett (Eds.), *The handbook of clinical intervention with young people who sexually abuse* (pp. 191–202). Hove, East Sussex, England: Brunner-Routledge.

Becker, J. V., Hunter, J., Stein, R., & Kaplan, M. S. (1989). Factors associated with erectile response in adolescent sex offenders. *Journal of Psychopathology and Behavioral Assessment*, *11,* 353–362.

Becker, J. V., Kaplan, M. S., & Tenke, C. E. (1992). The relationship of abuse history, denial and erectile response: profiles of adolescent sexual perpetrators. *Behavior Therapy*, *23,* 87–97.

Daleiden, E. L., Kaufman, K. L., Hilliker, D. R., & O'Neil, J. N. (1998). The sexual histories and fantasies of youthful males: A comparison of sexual offending, nonsexual offending, and nonoffending groups. *Sexual Abuse: A Journal of Research and Treatment*, *10,* 195–209.

Day, D. M., Miner, M. H., Sturgeon, V. H., & Murphy, J. (1989). Assessment of sexual arousal by means of physiological and self-report measures. In D. R. Laws (Ed.), *Relapse prevention with sex offenders* (pp. 115–123). New York: Guilford.

Glasgow, D. V. (2001). *Introduction to Affinity*. Unpublished document.

Glasgow, D. V., Osborne, A., & Croxen, J. (2003). An assessment tool for investigating paedophile sexual interest using viewing time: An application of single case methodology. *British Journal of Learning Disabilities*, *31,* 96–102.

Gray, S. R., & Plaud, J. J. (2005). A comparison of the Abel Assessment for Sexual Interest and penile plethysmography in an outpatient sample of sexual offenders. *Journal of Sexual Offender Civil Commitment: Science and the Law*, *1,* 1–10.

Gretton, H. M., McBride, M., Hare, R. D., O'Shaughnessy, R., & Kumka, G. (2001). Psychopathy and recidivism in adolescent sex offenders. *Criminal Justice and Behavior*, *28,* 427–449.

Gretton, H. M., Catchpole, R. E. H., McBride, M., Hare, R. D., O'Shaughnessy, R., & Regan, K. V. (2005). The relationship between psychopathy, treatment completion and criminal outcome over ten years: A study of adolescent sex offenders. In M. C. Calder (Ed.), *Children and young people who sexually abuse: New theory, research and practice developments* (pp. 19–31). Dorset, UK: Russell House Publishing.

Hanson, R. K., & Bussière, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology*, *66,* 348–362.

Hanson, R. K., & Morton-Bourgon, K. (2004). *Predictors of sexual recidivism: An updated meta-analysis (2004–2002)*. Ottawa, Canada: Public Safety and Emergency Preparedness Canada.

Harris, G. T., Rice, M. E., Quinsey, V. L., & Chaplin, T. C. (1996). Viewing time as a measure of sexual interest among child molesters and normal heterosexual men. *Behavior Research and Therapy*, *34,* 389–394.

Hunter, J. A., Goodwin, D. W., & Becker, J. V. (1994). The relationship between phallometrically measured deviant sexual arousal and clinical characteristics in juvenile sexual offenders. *Behavior Research and Therapy*, *32,* 533–538.

Hunter, J. A., & Lexier, L. J. (1998). Ethical and legal issues in the assessment and treatment of juvenile sex offenders. *Child Maltreatment*, *3,* 339–348.

Kaemingk, K. L., Koselka, M., Becker, J. V., & Kaplan, M. S. (1995). Age and adolescent sexual offender arousal. *Sexual Abuse: A Journal of Research and Treatment*, *7,* 249–257.

Konopasky, R. J., & Konopasky, A. B. (2000). Remaking penile plethysmography. In D. R. Laws, S. M. Hudson, & T. Ward (Eds.), *Remaking relapse prevention with sex offenders: A sourcebook* (pp. 257–284). Thousand Oaks, CA: Sage.

Lalumière, M. L., & Harris, G. T. (1998). Common questions regarding the use of phallometric testing with sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, *10,* 227–237.

Lang, A. R., Searles, J., Lauerman, R., & Adesso, V. (1980). Expectancy, alcohol, and sex guilt as determinants of interest in and reaction to sexual stimuli. *Journal of Abnormal Psychology*, *89,* 644–653.

Laws, D. R., Hanson, R. K., Osborn, C. A., & Greenbaum, P. E. (2000). Classification of child molesters by plethysmographic assessment of sexual arousal and a self-report measure of sexual preference. *Journal of Interpersonal Violence*, *15,* 1297–1312.

Letourneau, E. J. (2002). A comparison of objective measures of sexual arousal and interest: Visual reaction time and penile plyethysmography. *Sexual Abuse: A Journal of Research and Treatment*, *14,* 207–223.

Marshall, W. L., & Fernandez, Y. M. (2000). Phallometric testing with sexual offenders: Limits to its value. *Clinical Psychology Review*, *20,* 807–822.

McGrath, R. J., Cumming, G. F., & Burchard, B. L. (2003). *Current practices and trends in sexual abuser management: The Safer Society 2002 nationwide survey*. Brandon, VT: Safer Society Foundation.

Murphy, W. D., DiLillo, D., Haynes, M. R., & Steere, E. (2001). An exploration of factors related to deviant sexual arousal among juvenile sex offenders. *Sexual Abuse: A Journal of Research and Treatment*, *13,* 91–103.

Quinsey, V. L., Ketsetzis, M., Earls, C., & Karamanoukian, A. (1996). Viewing time as a measure of sexual interest. *Ethology and Sociobiology*, *17,* 341–354.

Seto, M. C. (2001). The value of phallometry in the assessment of male sex offenders. *Journal of Forensic Psychology Practice*, *1,* 65–75.

Seto, M. C., Lalumière, M. L., & Blanchard, R. (2000). The discriminative validity of a phallometric test for pedophilic interests among adolescent sex offenders against children. *Psychological Assessment*, *12,* 319–327.

Seto, M. C., Murphy, W. D., Page, J., & Ennis, L. (2003). Detecting anomalous sexual interests in juvenile sex offenders. *Annals of the New York Academy of Sciences*, *989,* 118–130.

Smith, G., & Fischer, L. (1999). Assessment of juvenile sexual offenders: Reliability and validity of the Abel assessment for interest in paraphilias. *Sexual Abuse: A Journal of Research and Treatment*, *11,* 207–216.

Worling, J. R. (2004). The Estimate of Risk of Adolescent Sexual Offense Recidivism (ERASOR): Preliminary psychometric data. *Sexual Abuse: A Journal of Research and Treatment*, *16,* 235–254.

Worling, J. R. (1998). Adolescent sexual offender treatment at the SAFE-T Program. In W. L. Marshall, Y. M. Fernandez, S. M. Hudson, & T. Ward (Eds.), *Sourcebook of treatment programs for sexual offenders* (pp. 353–365). New York: Plenum Press.